Rubric for Evaluating Extremely Small Schools
Under AZ LEARNS

Arizona State Board of Education
Information Packet
August 30, 2004

On August 30, 2004, the Arizona Department of Education (ADE) will present to the Arizona State Board of Education (Board) a proposed methodology for calculating an AZ LEARNS achievement profile for extremely small schools. This methodology will be applied to Arizona public elementary and secondary schools, including charter schools, classified as extremely small schools in order to determine school classifications by October 15, 2004 as required by A.R.S. §15-241 (Arizona LEARNS). The purpose of this document is to describe the proposed evaluation methodology to the Board.

Extremely small schools are defined as schools in which more than one-third of the AIMS subject-grade combinations cannot be evaluated due to having too few test scores in the baseline years 2000-2001. When AZ Learns achievement profiles were first issued in 2002 the Board determined that extremely small schools would not receive an achievement profile using the conventional AZ LEARNS methodology, and that ADE should develop an alternate method for evaluating these schools.

As mandated by A.R.S. §15-241, the ADE in collaboration with members of the education community developed the proposed evaluation methodology according to research-based principles. Upon adoption by the Board, the ADE will produce a technical report detailing the achievement profile methodology, including specific formulas and supporting documentation.

I. GENERAL PROCESS TO PRODUCE THE ACHIEVEMENT PROFILES

The method for calculating an achievement profile for extremely small schools is as follows:

A. AIMS scale score points will be calculated using a baseline score only. The baseline is calculated by aggregating test scores backward across years starting from the most current year. Aggregation is carried out until a group size of 32 is attained or the year 2000, whichever comes first. The percent of students passing AIMS in this aggregate group is then compared to the baseline grouping scale for the relevant subject/grade. The group then receives baseline points based upon which baseline group it is in.

B. Added evidence points will be calculated as for all other schools. Exception: Schools with less than 16 students in the MAP analysis will not receive added

evidence points. Instead, their total scale score points will be evaluated against a separate scale to determine a label.

C. Graduation/dropout points will be calculated as for all other schools.

D. The application of the percent exceeding thresholds for highly performing and excelling schools will be calculated as for all other schools.

E. The point scales for school labels will be the same as for all other schools.

F. Because of the uncertainty of measurement associated with small sample sizes and the high stakes of school labels, schools initially determined to be underperforming will receive a "second look." Instead of determining baseline groups based on the mean percent of students passing AIMS, an alternate baseline group for these schools will be determined based on the upper bound of a 95 percent confidence interval around the mean. If a school initially determined to be underperforming moves to a higher classification due to the "second look," that school will receive a "performing" label.

II. DEFINITION OF AN EXTREMELY SMALL SCHOOL

An extremely small school is defined as a school in which more than one-third of its subject/grade combinations cannot be evaluated because the average number of usable test scores in the baseline years of 2000-2001 is below the minimum group size (N-size or N-count) of 16.

Example.

The following table shows the number of usable test scores over the past four years for hypothetical middle school serving only eighth grade. For a test score to be usable it must be psychometrically valid and for a non-mobile student.

Table 1. An Extremely Small School

Subject	**2000**	**2001**	2002	2003	2004
Math	**15**	**15**	12	15	20
Reading	**15**	**17**	12	15	20
Writing	**14**	**14**	12	15	20

The minimum group size condition is applied by looking at the average group size in the baseline years of 2000-2001. If the average group size is below 16 then that subject/grade combination is not evaluated using the conventional AZ LEARNS methodology. In practice, the number of usable test scores in the baseline years must add up to 32 or greater. By this condition, reading can be evaluated (15 + 17 = 32), but writing and math cannot. Since more than one-third of the subject grade combinations

for this school cannot be evaluated, this school meets the definition of an extremely small school.

NOTES.

1. Group size for non-baseline years is not taken into account when determining if a school is extremely small.

2. State law (ARS 15-241) defines a small school as having less than 100 students. In the past a school must have had an average daily membership (ADM) of less than 100 to be considered extremely small. However, there are a significant number of schools with an ADM of greater than 100 yet fail to make the minimum group size standard. The evaluation method outlined here will be applied to all schools in which more than one-third of their subject/grade combinations cannot be evaluated, regardless of the schools' ADM.

III. CALCULATION OF AIMS SCALE SCORE POINTS

Subject/grade combinations with less than 16 students in the baseline years cannot be evaluated using the conventional AZLEARNS method. The baseline-growth method for calculating scale score points is too sensitive to changes in the scores of individual students for very small groups. Consequently, an alternate method has been developed to calculate AIMS scale score points for the AZLEARNS achievement profile.

The method is as follows:

A. For each subject/grade combination, an evaluation group is created by aggregating usable test scores across years starting in the most current year. Scores are added to the group until: 1) A group size of 32 is reached, or 2) Data for all available years is reached.

B. For each evaluation group, the percent of students who passed AIMS is calculated.

C. The percent passing for each evaluation group is compared to the AZLEARNS cutpoints for baseline groups.

D. Each subject/grade combination is awarded scale score points based on the baseline group it is in.

Example.

Table 2 below shows the number of usable tests for third grade reading for two different schools.

Table 2. Number of Usable Third Grade Reading Tests for Two Schools

School	2000	2001	2002	2003	2004	Total Group Size
A	14	15	**13**	**12**	**15**	40
B	**4**	**5**	**6**	**6**	**5**	26

Applying the methodology described above, for each school we form an evaluation group by counting backward from the most current year until either we achieve a group size of 32 or use all the available data. The years of data used for each school are shown in **boldface.** For school A this results in a group of 40 students from the years 2002, 2003 & 2004. For school B, we must use all the years of available data back to 2000.

The next step is to examine the percent of students passing in the groups we have formed. Table 3 shows the number of students who passed in the groups from Table 2. Again for school A we consider only the most recent three years while for school B we consider five years.

Table 3. Number of Students Passing AIMS, Third Grade Reading

School		2000	2001	2002	2003	2004	Total	% Pass
A	# Pass	NA	NA	8	8	5	21	53
	# Tested	NA	NA	**13**	**12**	**15**	40	
B	# Pass	2	3	5	5	5	20	77
	# Tested	**4**	**5**	**6**	**6**	**5**	26	

For school A, 53 percent of the students in the evaluation group have passed AIMS. Comparing this to the cutpoints for baseline groups approved last year by the Board (Table 4) we see that this puts third grade reading for school A in baseline grouping two, earning the school two scale score points.

For school B, 77 percent of students in the evaluation group have passed AIMS. Comparing this to the cutpoints for baseline groups (Table 4) we see that this puts third grade reading for school B in baseline grouping four, earning the school four scale score points.

Table 4. Baseline Groupings

Grade	Subject	Baseline Grouping 1	Baseline Grouping 2	Baseline Grouping 3	Baseline Grouping 4	Baseline Grouping 5	Baseline Grouping 6
3	Math	0% - 26%	27% - 40%	41% - 56%	57% - 71%	72% - 82%	83% - 100%
3	Reading	0% - 46%	47% - 59%	60% - 73%	74% - 84%	85% - 91%	92% - 100%
3	Writing	0% - 54%	55% - 67%	68% - 79%	80% - 89%	90% - 94%	95% - 100%
5	Math	0% - 11%	12% - 21%	22% - 36%	37% - 52%	53% - 66%	67% - 100%
5	Reading	0% - 31%	32% - 44%	45% - 60%	61% - 75%	76% - 85%	86% - 100%
5	Writing	0% - 25%	26% - 38%	39% - 53%	54% - 68%	69% - 79%	80% - 100%
8	Math	0% - 1%	2% - 5%	6% - 12%	13% - 22%	23% - 34%	35% - 100%
8	Reading	0% - 25%	26% - 37%	38% - 51%	52% - 66%	67% - 77%	78% - 100%
8	Writing	0% - 18%	19% - 28%	29% - 42%	43% - 56%	57% - 68%	69% - 100%
H.S.	Math	0% - 3%	4% - 8%	9% - 19%	20% - 33%	34% - 47%	48% - 100%
H.S.	Reading	0% - 28%	29% - 42%	43% - 58%	59% - 73%	74% - 83%	84% - 100%
H.S.	Writing	0% - 16%	17% - 25%	26% - 39%	40% - 53%	54% - 66%	67% - 100%

BOARD ACTION:

(A) The ADE recommends that the Board adopt the methodology for calculating the AIMS scale score points for extremely small schools described above.

IV. CALCULATION OF AN ACHIEVEMENT PROFILE FOR AN EXTREMELY SMALL SCHOOL

In addition to AIMS scale score points, schools also earn scale score points via added evidence, their graduation and dropout rates, and their current AYP status. The methods used to calculate scale score points earned by extremely small schools for these performance measures will be the same as the methods used for other schools approved by the Board on September 16, 2003.

Figures A and B give a graphical summary of the method for evaluating extremely small schools at the elementary and high school levels.

- A school may receive up to six scale score points (calculated using the method described in III) for each subject/grade combination it serves. This is represented by the far left column in each figure.

- A school may receive one point if it has made adequately yearly progress (AYP) under the methodology mandated by the No Child Left Behind Act. This is represented by the column second-from-left in each figure.

- If a school is an elementary school, it may receive added evidence points based on it performance on the Measure of Academic Progress (MAP). This is the fourth-from-left column in Figure A.

- If a school is a high school it may receive up to two points based on its graduation and dropout rates. This is the third-from-left column in Figure B.

- The total points earned by a school are added up and compared to the school classification scale to determine a school's preliminary classification—the final column in Figures A and B.

- In order to be classified as a "highly performing" or "excelling" school, a school must meet the thresholds for percentage of students exceeding the standard approved by the Board on September 16, 2003.

Figure A. Method for Evaluating Extremely Small Schools (Elementary)

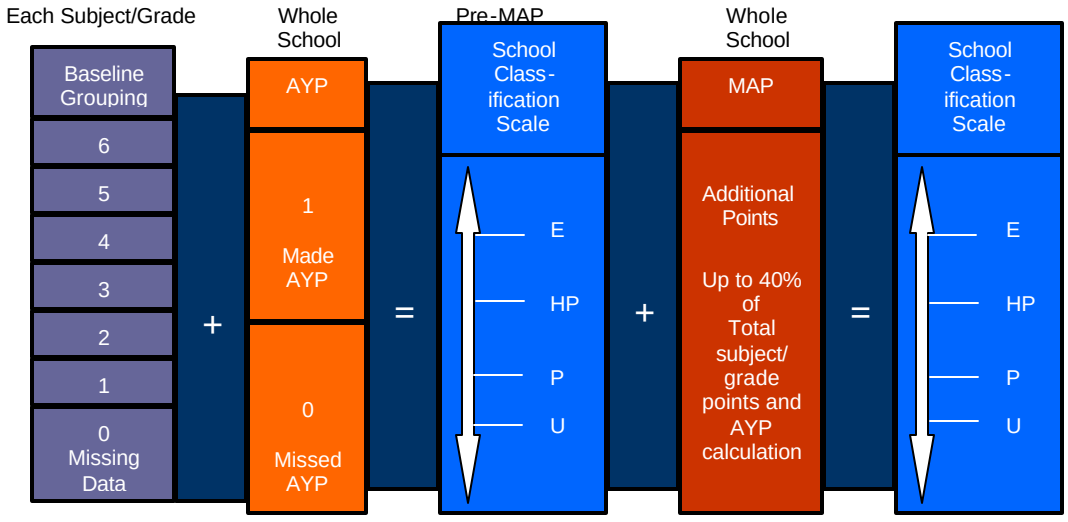
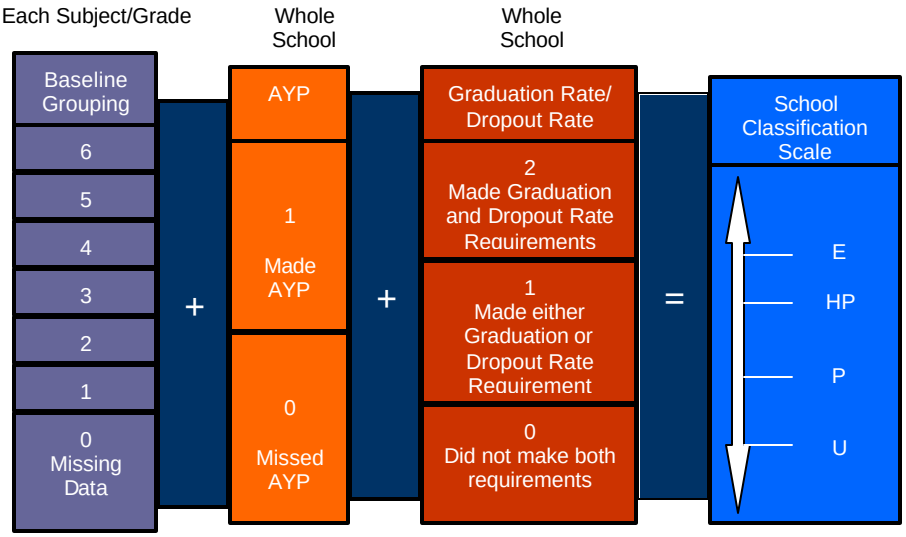


Figure B. Method for Evaluating Extremely Small Schools (High Schools)



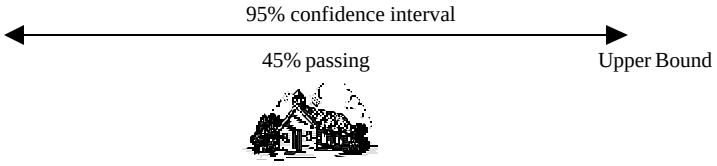
- **<u>"Second look" for extremely small schools.</u>** Because of the high-stakes consequences of being labeled an "underperforming" school, and because of the uncertainty of measurement involved with small sample sizes, the ADE believes it is prudent to give extremely small schools a "second look" if they face the possibility of receiving an "underperforming" label. If the preliminary label of an extremely small school is "underperforming," then the AIMS scale score points for that school will be recalculated. For each subject/grade combination, the upper bound of the 95-percent confidence interval will be used to calculate to which baseline group the school belongs. If the recalculated points move the school into a higher classification, the school will receive a "performing" label.

<u>Example.</u>

Figure C shows how the "second look" will work. Forty-five percent of the students in a hypothetical school have passed third grade math. This places the school in baseline grouping three for this subject and grade, and earns the school three AIMS scale score points. If this school is "underperforming" then a 95-percent confidence interval is calculated using the normal approximation to the binomial distribution. As Figure C shows the upper bound of this confidence interval places the school in baseline grouping five—earning the school five AIMS scale score points instead of three. This recalculation is performed for every subject/grade combination served by the school. If the increase in scale score points is sufficient, the school will receive a "performing" label.

Figure C. Example of a Second Look for an Underperforming School

Baseline Grouping 1	Baseline Grouping 2	Baseline Grouping 3	Baseline Grouping 4	Baseline Grouping 5	Baseline Grouping 6
0% - 26%	27% - 40%	41% - 56%	57% - 71%	72% - 82%	83% - 100%



BOARD ACTION:

(B) The ADE recommends that for the extremely small school achievement profile the Board adopt the calculation of scale score points for added evidence (MAP), graduation and dropout rates, adequate yearly progress (AYP), and the application of the percent-exceeding thresholds for "highly performing" and "excelling" schools using the same methods and parameters approved September 16, 2003.

(C) The ADE recommends that the Board adopt the recalculation of AIMS scale score points for underperforming extremely small schools described above.

(D) The ADE recommends that the Board adopt the policy that if the recalculation of AIMS scale score points using the methodology described above results in sufficient scale score points for a school to not be labeled "underperforming" that the school receive a "performing" label.

V. CLASSIFICATION SCALES FOR EXTREMELY SMALL SCHOOLS

ADE proposes that the classification scales used for extremely small schools be the same scales used for other schools approved by the Board on September 16, 2003 (shown below). The ADE also proposes that if an extremely small school has less than sixteen students in its MAP analysis, the school should not receive added evidence points and instead be classified using the scale in Table 6 It is likely that extremely small schools would have very few or no students matched across years for the MAP analysis. Consequently, the MAP analysis, and thus the points earned for an achievement profile, would be sensitive to the performance of one or two students. In order to avoid this problem, ADE is proposing that elementary schools with less than sixteen students in the MAP analysis not receive added evidence points and be evaluated against a separate scale.

Table 5. Elementary School Classification Cut Points

	Subject/Grade Combination 1	Subject/Grade Combination 3	Subject/Grade Combination 6	Subject/Grade Combination 9
Underperforming	< 4	< 12	< 24	< 36
Performing	4	12	24	36
Highly Performing	4.6	13.8	27.6	41.4
Excelling	5.4	16.2	32.4	48.6

Table 6. High School Classification/Non-MAP Elementary School Cut Points

	Subject/Grade Combination 1	Subject/Grade Combination 3
Underperforming	< 3.2	< 9.6
Performing	3.2	9.6
Highly Performing	5	15
Excelling	5.4	16.2

BOARD ACTION:

(E) The ADE recommends that the Board adopt the school classification scales given in Tables 5 and 6 for the evaluation of extremely small schools.

(F) The ADE recommends that the Board adopt the policy that extremely small schools with less than 16 students in the MAP analysis do not receive added evidence points and be measured against the classification scale in Table 6.